

# **GMG: A Guaranteed, Efficient Global Optimization Algorithm for Remote Sensing**

April 2004

Prepared by  
C. D'Helon  
V. Protopopescu  
J. C. Wells  
J. Barhen

#### DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge:

**Web Site:** <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
**Telephone:** 703-605-6000 (1-800-553-6847)  
**TDD:** 703-487-4639  
**Fax:** 703-605-6900  
**E-mail:** [info@ntis.fedworld.gov](mailto:info@ntis.fedworld.gov)  
**Web site:** <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE), and International Nuclear Information System (INIS) representatives from the following sources:

Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831  
**Telephone:** 865-576-8401  
**Fax:** 865-576-5728  
**E-mail:** [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)  
**Web site:** <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**GMG: A GUARANTEED, EFFICIENT GLOBAL OPTIMIZATION  
ALGORITHM FOR REMOTE SENSING**

C. D'Helon  
V. Protopopescu  
J. C. Wells  
J. Barhen

Date Published: April 2004

Prepared by  
OAK RIDGE NATIONAL LABORATORY  
P. O. Box 2008  
Oak Ridge, Tennessee 37831-6285  
managed by  
UT-Battelle, LLC  
for the  
U. S. DEPARTMENT OF ENERGY  
under contract DE-AC05-00OR22725



# Contents

<b>LIST OF FIGURES</b>	<b>v</b>
<b>ACKNOWLEDGEMENTS</b>	<b>vii</b>
<b>ABSTRACT</b>	<b>ix</b>
<b>1 INTRODUCTION - THE GLOBAL OPTIMIZATION PROBLEM</b>	<b>1</b>
<b>2 MONOCULAR PASSIVE RANGING</b>	<b>3</b>
<b>3 A GLOBAL OPTIMIZATION ALGORITHM WITH A GUARANTEE: GMG</b>	<b>7</b>
3.1 SUFFICIENT ADDITIONAL INFORMATION . . . . .	7
3.2 INCOMPLETE ADDITIONAL INFORMATION . . . . .	8
<b>4 GMG ASSESSMENT ON BENCHMARK FUNCTIONS</b>	<b>13</b>
<b>5 SOLVING THE MPR PROBLEM</b>	<b>21</b>
5.1 GMG ASSESSMENT ON THE MPR PROBLEM . . . . .	24
<b>6 CONCLUSION</b>	<b>25</b>
<b>BIBLIOGRAPHY</b>	<b>27</b>



## List of Figures

1	Baseline uplooking MPR scenario for an airborne sensor. . . . .	4
2	Total processor time to find the global minimum with a guarantee vs. the number of dimensions for exactly-known ( $\beta = \beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta = \beta_1 = 0.1, \beta_2 = 0.2$ ; hollow dots) golf-course functions. . . .	9
3	Total processor time to find the global minimum with a guarantee vs. the grid size for exactly-known ( $\beta_0 \geq \beta$ ; solid dots) and “uncertain” ( $\beta_1 \geq \beta$ ; hollow dots) two-dimensional golf-course functions. . . . .	9
4	Probability to find the global minimum vs. the grid size for exactly-known ( $\beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta_1 = 0.1, \beta_2 = 0.2$ ; hollow dots) one-dimensional golf-course functions. . . . .	14
5	Probability to find the global minimum vs. a fixed running time, for exactly-known ( $\beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta_1 = 0.1, \beta_2 = 0.2$ ; hollow dots) one-dimensional golf-course functions. . . . .	14
6	The Branin objective function. . . . .	17
7	The Camelback objective function. . . . .	17
8	The Goldstein-Price objective function. . . . .	18
9	The Rastrigin objective function. . . . .	18
10	The Shubert objective function. . . . .	19
11	The Hartman objective function. . . . .	19
12	Standard mean ozone profile for mid-latitudes in summer. The ozone density unit used in the graph is the Dobson Unit ( $1 \text{ DU} = 2.14 \times 10^{-8} \text{ kg/m}^3$ ). . .	21
13	Error function using three ozone layers, for (a) mid-ozone concentration change ( $\pm 1\%$ ) vs. range (km), and b) low-ozone concentration change ( $\pm 1\%$ ) vs. range (km). The error function is plotted on a logarithmic scale, with dark intensities indicating smaller values. The position of the global minimum is indicated by the superimposed arrows. . . . .	23





## ACKNOWLEDGEMENTS

We thank Drs. David Reister and Neena Imam for their useful suggestions and assistance. Our research was sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL) (CDH, VP, JCW), and by the Division of Materials Sciences and Engineering, U.S. Department of Energy (US DOE) (VP, JCW, JB), and used the resources of the Center for Computational Sciences at ORNL supported by the Office of Science, US DOE, all under contract No. DE-AC05-00OR22725 with UT-Battelle, LLC.



## ABSTRACT

The monocular passive ranging (MPR) problem in remote sensing consists of identifying the precise range of an airborne target (missile, plane, etc.) from its observed radiance. This inverse problem may be set as a global optimization problem (GOP) whereby the difference between the observed and model predicted radiances is minimized over the possible ranges *and* atmospheric conditions. Using additional information about the error function between the predicted and observed radiances of the target, we developed GMG, a new algorithm to find the Global Minimum with a Guarantee. The new algorithm transforms the original continuous GOP into a discrete search problem, thereby *guaranteeing* to find the position of the global minimum in a reasonably short time. The algorithm is first applied to the golf course problem, which serves as a litmus test for its performance in the presence of both complete and degraded additional information. GMG is further assessed on a set of standard benchmark functions and then applied to various realizations of the MPR problem.



# 1 INTRODUCTION - THE GLOBAL OPTIMIZATION PROBLEM

The Global Optimization Problem (GOP) sounds deceptively simple: find the absolute minimum of a given function - called the objective function - over the allowed range of its variables.

The primary difficulty in solving GOPs stems from the fact that the condition for determining minima, namely annulment of the gradient of the objective function, is only *necessary* i.e., as it may indicate another type of critical point, and *local*, as it does not distinguish between local and global minima. Since, in general, the number of local minima grows exponentially with dimensionality, solving the GOP turns out to be a computationally hard problem [1].

The generic strategy to find the global minimum involves two main operations, namely: (i) descent to a local minimum, and (ii) search for a new descent region, which are alternatively repeated. Since each function evaluation may involve an expensive computational process, the number of function evaluations needs to be kept to a minimum. It is not surprising that, together with accuracy, this number provides the paramount criterion in comparing the efficiency of competing optimization algorithms. However, as the dimensionality of the problem increases, the search phase becomes the most time-consuming part of the algorithm. As a result, *in general*, conventional strategies offer no guarantee and little hope that the global minimum could be found in a reasonable time.

To guarantee finding the location of the global minimum while maintaining the search at an affordable cost, additional information about the objective function has to be provided and used. Traditionally this information has been mostly related to various smoothness properties of the objective function, such as global uniform bounds on its first and/or second partial derivatives with respect to the independent variables. If such information is available, it can be used to construct a domain covering which is then exhaustively or selectively searched. For a good review on covering methods, the interested reader is referred to the book of Törn and Žilinkas [2].

The method presented here belongs to the same general framework. Its novelty consists in the *type* of information that is used to reduce the complexity of the GOP. The relevance of our method is threefold: (i) it does not require information on smoothness, which often is difficult to obtain; (ii) whenever information on smoothness *is available*, our information complements it; and (iii) it ensures that the resulting covering is the minimal one needed to guarantee finding the global minimum.

Interest in developing efficient and reliable methods for the GOP has grown together with the number of its practical applications [3], which range from logistics and transportation to pattern recognition for biometric identification and protein folding in computational biology. In the following, we focus on the specific GOP that arises when solving the monocular passive ranging (MPR) problem [4, 5] for remote sensing in the atmosphere.



## 2 MONOCULAR PASSIVE RANGING

Remote sensing in the atmosphere is an important field of research that is directly relevant to National Defense. For example, a recent American Physical Society study investigates the challenges of boost-phase defense against intercontinental ballistic missiles [6], pointing out the importance of being able to detect, identify, and track missiles as early and as precisely as possible. Thus, solving the MPR efficiently and accurately represents the technical foundation for critical defense decisions (e.g., in missile defense), whence the importance of a global optimization algorithm that guarantees the optimal solution in a reasonable time.

The goal of MPR is to enable accurate target range estimation when active measurements (e.g., radar) or multiple views of the same target are unavailable. Therefore the instantaneous range to a target is determined using a single sensor. If the sensor is fitted on an aircraft operating at cruising altitude, then two measurement scenarios are possible: uplooking, when the target is at a higher altitude (see Fig. 1), and downlooking, when the target is at a lower altitude.

As electromagnetic radiation travels through the atmosphere, it interacts with molecules of various gases and larger particles, and is either transmitted, absorbed, or scattered. The result of the absorption and scattering processes is an effective extinction of the radiation at a rate that depends on wavelength, temperature, density, aerosol composition, and the absorption coefficient for each gas or particle, which in turn depends on the atomic and molecular structure. In general, the target range in MPR depends on the cumulative effect that many different atmospheric variables have on the propagation of the radiation detected by the sensor. The calculation of the radiance at the sensor is sensitive in a nonlinear way to various uncertainties. Therefore, determining the target range amounts to defining and solving a quite complex inverse problem that poses a formidable computational challenge.

Modeling the absorption and scattering processes to calculate the radiance of a target can be achieved using MODTRAN. MODERate resolution TRANsmission is a computational model [7, 8] developed by the Air Force Research Lab / Space Vehicles Directorate and Spectral Sciences Inc., for predicting atmospheric radiation transmission, radiance, and flux for targets at different ranges. It is an authoritative benchmark for calculating the radiance received at an airborne or spaceborne sensor, across a full spectral range, for arbitrary refracted paths above the curved earth, and allows the user to create rich descriptions of the atmosphere, including vertical profiles of the water vapor, ozone, aerosols, etc.

The computational techniques used in MPR have progressed from basic approximations [5] that make ad-hoc assumptions about atmospheric conditions and the target radiance to advanced methods [4], based on continuum interpolation or parameter estimation.

To date, the most sophisticated method for MPR [4] is parameter estimation, which uses an error function between the actual sensor-measured radiance of a target, and the radiance calculated by MODTRAN for different ranges and atmospheric conditions. The range is obtained by minimizing the error function in the least squares sense, starting from an educated guess for the initial values of the relevant parameters. The computational complexity of this multi-dimensional nonlinear optimization problem is simplified by linearizing the MODTRAN model, and retaining only the first term of its Taylor expansion, which is generated by an automatic differentiation process [9].

Our approach to solving the MPR also employs an optimization of the error function defined as the square of the difference between the actual observed radiance of a target and the radiance calculated by MODTRAN. However, here we consider the original problem (as opposed to its linearization) and we aim to *guarantee* finding the global minimum of the error

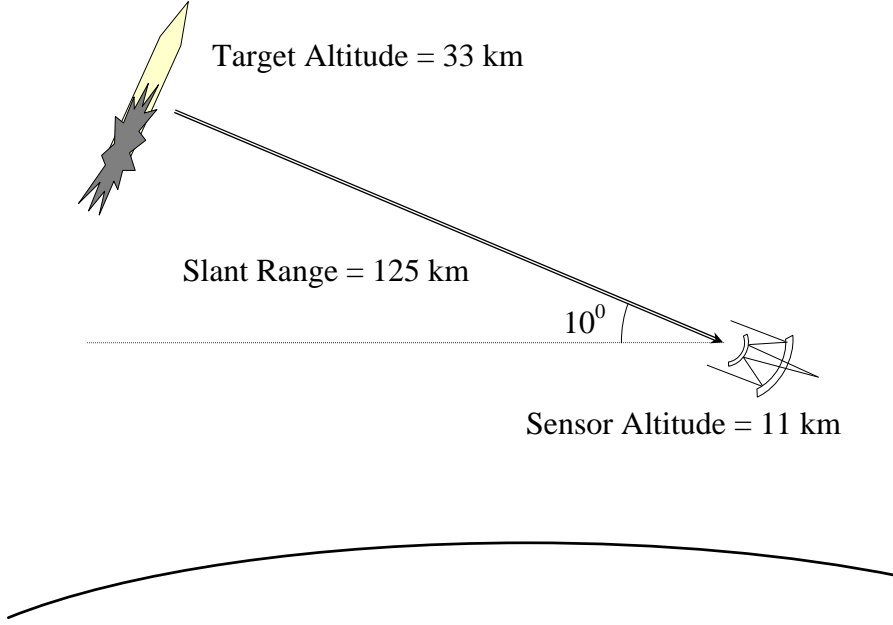


Figure 1: Baseline uplooking MPR scenario for an airborne sensor.

function. To illustrate our approach, we restrict ourselves to a *very simple* MPR problem, namely: determine the range from *one* sensor measurement, for density fluctuations of only *one* type of absorbing gas.

This MPR problem can be cast as a GOP using the following steps:

1. Obtain a sensor measurement,  $S$ , of the radiance of a target, which is at an (yet to be determined) range  $x_0$  from the sensor.
2. Find the global minimum of the error function,  $\mathcal{E}(\zeta)$ , between the observed radiance,  $S$ , and the model-calculated radiances,  $\mathcal{R}(\zeta)$ , over the entire domain  $\zeta$ ,

$$\min \mathcal{E}(\zeta) = \min ||S - \mathcal{R}(\zeta)||^2, \quad (1)$$

to determine the actual range  $x_0$ .

3. Use MODTRAN to calculate the radiances,  $\mathcal{R}(\zeta)$ , at the sensor, for various realizations of the "variable"  $\zeta = \{x, \rho(r)\}$ , where  $x$  is the range and  $\rho(r)$  is a function that depends on the atmospheric altitude,  $r$ , and is assumed to describe *all* the pertinent features of the atmosphere.

Since measured radiances are known to be most sensitive to fluctuations in the ozone density, we restrict ourselves to a situation where the function  $\rho(r)$  represents *only the ozone density*, at the altitude  $r$ . We note that, in general, the ozone profile depends on



various factors such as latitude, calendar month, proximity to built-up areas, etc. We can take advantage of measurements available for ozone profiles in different scenarios to further reduce the complexity of the problem. This additional information is used by considering small variations,  $\delta\rho(r)$ , about a known standard mean ozone profile,  $\rho_0(r)$ , in the form:

$$\rho(r) = \rho_0(r) + \delta\rho(r). \quad (2)$$

The variations  $\delta\rho(r)$  are functions that belong to an *infinite dimensional* set. We discretize  $\delta\rho(r)$  by  $k$  independent variables,  $\delta\rho_1, \delta\rho_2, \dots, \delta\rho_k$ , corresponding to  $k$  atmospheric layers,  $l_1, \dots, l_k$ . This results in an *approximate* GOP, whereby we minimize the error function:

$$E(\mathbf{y}) = \|S - R(x, \delta\rho_1, \delta\rho_2, \dots, \delta\rho_k)\|^2, \quad (3)$$

over a bounded  $(k+1)$ -dimensional vector  $\mathbf{y} = \{x, \delta\rho_i\}$ . Henceforth, we consider the solution to the problem in Eq. 3 to be “exact”, and therefore all further errors would correspond to deviations from the solution to this problem.

To assess the results of our algorithm, we do not use a measured radiance  $S$ . Instead, we use MODTRAN to compute a response for a given set of known parameter vector  $\mathbf{y}^* = \{x_{0S}, \delta\rho_{1S}, \delta\rho_{2S}, \dots, \delta\rho_{kS}\}$ . This response is used as  $S$ , and at the end of the optimization process we expect to recover the known parameter vector  $\mathbf{y}^*$ .

Indeed, when the parameters used in the model yield a radiance that coincides exactly with the measured radiance from the target, the global minimum is attained, since the value of the global minimum for a perfect fit is zero. Assuming that the optimal solution is unique, all the other minima of the error function, Eq. 3, have positive values.



### 3 A GLOBAL OPTIMIZATION ALGORITHM WITH A GUARANTEE: GMG

To guarantee finding the global minimum in a timely manner, additional information has to be used to reduce the complexity of the GOP. Additional information about the error function,  $E(\mathbf{y})$ , may include the size of the basin of attraction for the global minimum, the separation between the global minimum and the next lowest local minimum, the value of the global minimum (e.g., in the example above it is zero), the Lipschitz constants, the general shape of the error function, and information about the uncertainties inherent in the sensor measurement. We note that while some amount of information is *usually available* in many applications, existing optimization algorithms *do not* and sometimes *cannot exploit* this information. As a result, they cannot offer any guarantee that the global minimum has been found.

The additional information about the objective function enables us to map the continuous GOP onto a discrete search, whereby: (i) all basins of attraction, except that of the global minimum, are eliminated, and (ii) the number of function evaluations in the search process is kept at the minimum number necessary to guarantee the optimal solution. Below, we present a specific set of such conditions. Further on, we demonstrate that these assumptions may be relaxed, and consider the effect of incomplete information on the performance and guarantee of the algorithm.

#### 3.1 SUFFICIENT ADDITIONAL INFORMATION

Consider an error function in MPR for  $d = k + 1$  variables,  $E(\mathbf{y})$ , that has a certain degree of smoothness. Without restricting generality, we can assume that  $E$  is defined on the domain  $[0, 1]^d$  and takes values in the range  $[0, 1]$ . Different (bounded) domains and ranges can be accommodated by appropriate scalings.

The following conditions constitute sufficient additional information about the error function, for our algorithm to find the global minimum with a guarantee:

1. there is a unique global minimum and its value is zero,  $E_{min} = 0$ ;
2. there are no local minima whose value is infinitesimally close to zero; i.e., the values of the other minima are larger than a constant  $\delta > 0$ , and
3. the size of the basin of attraction for the global minimum, measured at height  $\delta$ , is known.

The threshold value  $\delta$  identifies the values of the error function that are smaller than  $\delta$  and belong exclusively to the basin of attraction of the global minimum, thereby enabling its identification.

The domain of the error function is discretized, by uniformly dividing  $[0, 1]^d$  into a grid of  $M^d$  small  $d$ -dimensional hypercubes, which provides a complete covering [2]. The information about the size of the basin of attraction of the global minimum specifies the necessary resolution of the discretization grid [15]. The error function is evaluated at  $M + 1$  points in each dimension, in order to guarantee that only one of the  $N = (M + 1)^d$  points belongs to the basin of attraction of the global minimum.

Then, the error function is shifted upwards by the known value  $1 - \delta$ , and its integer part,

$$e(\mathbf{y}_i) = \text{INT}[E(\mathbf{y}_i) + (1 - \delta)], \quad (4)$$

is evaluated at a series of points  $\mathbf{y}_i$  ( $i = 1 \dots N$ ).

Thus the range of  $e(\mathbf{y})$  is limited to the values zero and one, mapping the original continuous GOP onto a discrete, unsorted search problem. This transformed error function is known to be equal to one for all inputs, except for  $\mathbf{y} = \mathbf{y}^*$ , where  $e(\mathbf{y}^*) = 0$ . We note that searching for the value  $\mathbf{y}^*$ , which identifies the basin of attraction of the global minimum, is equivalent to searching for the golf-hole in the notoriously difficult golf-course problem (see below). Once  $\mathbf{y}^*$  is found, an appropriate descent method is applied to the original function to reach the global minimum, and consequently determine the exact range,  $x_0$ , of the target from the sensor.

Thus finding the Global Minimum is Guaranteed in a finite time, as long as the number of function calls used in the discrete search phase is finite. For this reason we call our algorithm GMG. For a small number  $N$  of error function evaluations, there are simple and straightforward algorithms for the discrete, unsorted search that can be run on a single processor in a reasonable time. However, for GOPs that require high-density grids, or have a large number of dimensions, this search has to be implemented on high-performance computers. Parallel algorithms for the discrete, unsorted search are “embarrassingly parallel”, thus they can utilize the full power of available machines with large numbers of processors.

Further reductions in complexity are offered by quantum computing [10], or an exponentially large number of processors [11]. We note that while the original curse of dimensionality is not eliminated by our algorithm, it is nevertheless mitigated by bringing the number of function evaluations to the *absolute minimum* necessary to *guarantee* the optimal result.

To validate the algorithm we applied it to the golf-course problem, which offers all the additional information listed above. The one-dimensional version of the golf-course problem can be written as:

$$f(x) = \begin{cases} 0 & \text{for } a - \beta_0/2 \leq x \leq a + \beta_0/2 \\ 1 & \text{for } 0 \leq x \leq a - \beta_0/2 \text{ and } a + \beta_0/2 \leq x \leq 1 \end{cases} , \quad (5)$$

where  $a$  is any point in the interval  $(\beta_0/2, 1 - \beta_0/2)$ , but is otherwise unknown.

For this function, the size  $\beta_0$  of the basin of attraction is known exactly. To ensure that there is a unique global minimum located in the interval  $a - \beta_0/2 \leq x \leq a + \beta_0/2$ , the grid size for the discretization of the domain of  $f(x)$  is set to  $\beta = \beta_0$ . Thus, for the golf-course problem, all of the sufficient additional information discussed in Section 3.1 is available, and finding the global minimum with a guarantee is equivalent to the discrete search problem for the error function  $e(\mathbf{y})$  in Eq. (4).

The  $d$ -dimensional version of the golf-course problem can be written as:

$$f(x_i) = \begin{cases} 0 & \text{for } a_i - \beta_0/2 \leq x_i \leq a_i + \beta_0/2, \text{ for all } i = 1 \dots d \\ 1 & \text{otherwise} \end{cases} , \quad (6)$$

where  $a_1, \dots, a_d$  are the coordinates of the center of the  $d$ -dimensional hypercubic golf-hole.

### 3.2 INCOMPLETE ADDITIONAL INFORMATION

The set of conditions listed in the previous section illustrate the type and amount of additional information that is sufficient to guarantee that our algorithm will find the global minimum. However, in general, these conditions are satisfied only to a certain extent, due both to incomplete knowledge and noise inherent in measurements. Missing or corrupted information results in the following relaxations of the sufficient conditions:

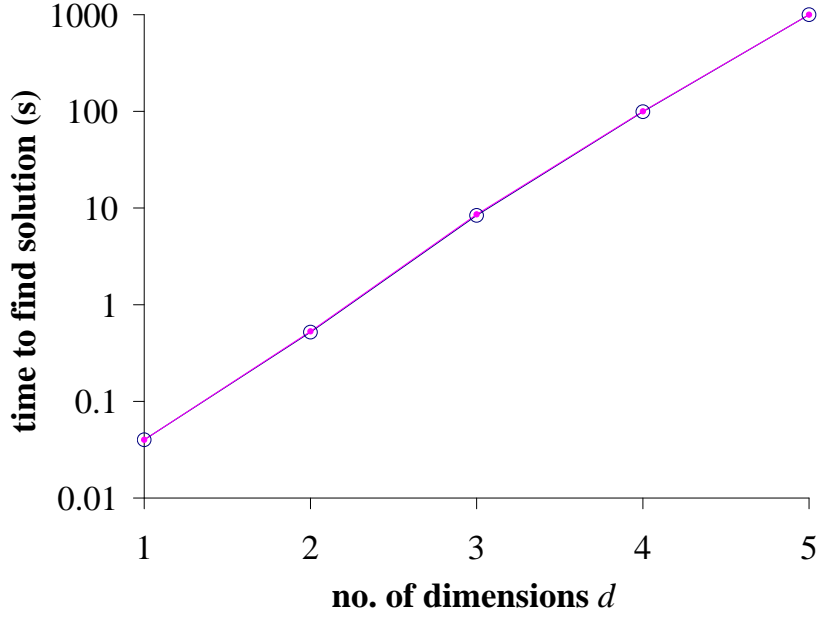


Figure 2: Total processor time to find the global minimum with a guarantee vs. the number of dimensions for exactly-known ( $\beta = \beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta = \beta_1 = 0.1$ ,  $\beta_2 = 0.2$ ; hollow dots) golf-course functions.

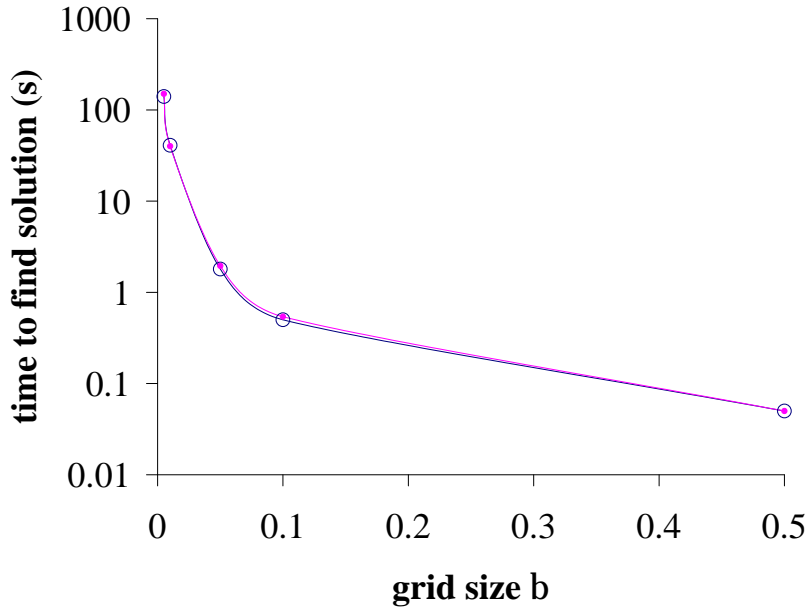


Figure 3: Total processor time to find the global minimum with a guarantee vs. the grid size for exactly-known ( $\beta_0 \geq \beta$ ; solid dots) and “uncertain” ( $\beta_1 \geq \beta$ ; hollow dots) two-dimensional golf-course functions.

1. Theoretically, the error function  $E(\mathbf{y})$  should be zero when the predicted radiance is exactly equal to the observed target radiance. However, the global minimum of the error function may have a value greater than zero, due to measurement noise.
2. The value  $\delta$  separating the global minimum from the next-lowest minimum may be inaccurate or completely unknown. Still, if the total number of minima is finite, which is a very reasonable assumption for a real-world problem, then such a finite  $\delta$  does exist.
3. Finally, we often have only a rough estimate or an upper bound for the size of the basin of attraction of the global minimum. Of course, one may attempt to estimate this size from previous samplings of  $E(\mathbf{y})$ , whenever available.

We report the performance of GMG under the effect of insufficient and/or corrupted information, as described in the conditions 1-3 above.

1. The global minimum of the error function can still be found after the addition of noise as long as it retains its "global minimum" status. Typically, this means that in solving the MPR problem, we can guarantee to find the noise-free global minimum if the noise level is relatively small ( $\leq \delta$ ). In general, if the ordering of the  $j$  lowest minima in the error function is scrambled, our algorithm can find all of these minima, but then offers no help in sorting them out correctly using only a *single* measurement.

2. Finding the noise-free global minimum is straightforward if we know the value of  $\delta$ . Otherwise, we have to iterate the transformation in Eq. 4 for the error function. We start with  $\delta = 1/2$ , which is likely to produce many zero values in the transformed error function, and keep refining  $\delta$  by a halving procedure until exactly one zero value is found in the output of  $e(\mathbf{y})$ . The corresponding point is guaranteed to be in the basin of attraction of the global minimum, assuming that the size of the basin of attraction is known.

The estimation of the threshold  $\delta$  requires no additional evaluations of the error function, and only a logarithmic number of iterative steps  $O(\log\{1/\delta\})$  to identify the basin of attraction. Thus our algorithm can deal successfully with these types of missing information, at the same time, given that the ordering of the minima is unchanged in the presence of noise.

3. Information about the size of the basin of attraction is more consequential, since it controls the minimum grid density, i.e., the required number of error function evaluations, which can be expensive. This also explains why we do not start with a very small value of  $\delta$ , which would entail a grid density greater than the required minimum. The discretization grid in Eq. 4 has to be sufficiently dense though to ensure that the global minimum maintains its absolute minimum status in the process of sampling the error function. If the function evaluations required by the grid can be computed in a reasonable time, then it is straightforward to iterate the estimate of the shift  $1 - \delta$  for the error function, until exactly one zero value is found in the output of  $e(\mathbf{y})$ .

For the simplest case of a one-dimensional objective function, where the smallest grid size available for a fixed running time is  $\beta$ , and the actual size of the basin of attraction is  $\beta_0$ , our algorithm's outcomes can be summarized as follows:

1. if  $\beta_0$  is known exactly, and
  - (a)  $\beta \leq \beta_0$ : the algorithm is guaranteed to find the global minimum;
  - (b)  $\beta \gg \beta_0$ : the algorithm has only a small chance to find the global minimum;

2. if  $\beta_0$  is completely unknown: use  $\beta$  for the grid size, and
  - (a) if  $\beta \leq \beta_0$ : the algorithm will find the global minimum;
  - (b) else the algorithm will most likely NOT find the global minimum;
3. if a probability distribution has been calculated, measured, or inferred for  $\beta_0$  (e.g., an upper bound) then the cumulative probability for  $\beta_0 > \beta$  determines how likely it is that the algorithm will find the global minimum.

The performance of GMG under these relaxed conditions has been again verified on suitably modified golf-course problems. We assume that the “uncertain” golf-course function maintains the same shape, however the size of the basin of attraction varies from  $\beta_1$  to  $\beta_2$ , within a well-defined uniform probability distribution.

The results of our numerical tests for the golf-course problem with precise and imprecise additional information are summarized in Figures 2-5. Figures 2 and 3 show a comparison of the time required to find the global minimum with a guarantee, after averaging over a hundred GOP searches with random values of  $a$ . The results for both cases are almost identical, since  $\beta_0$  and  $\beta_1$  have been set equal for the comparison. Figures 4 and 5 show a comparison of the probability to find the global minimum, for averages over a thousand GOP searches with random values of  $a$ .

Both types of golf-course functions lead to the same conclusions about the performance and time scaling of our algorithm:

1. The time taken to find a solution grows exponentially with the number of dimensions - this is a direct and, in general, unavoidable consequence of the dimensionality curse.
2. The time taken to find a solution grows as an inverse power of the size of the basin of attraction  $t \propto \beta_0^{-d}$ , for a fixed number of dimensions  $d$ . As expected, the denser the grid, the more error function evaluations are required.
3. Finding the global minimum is certain if the minimum size of the basin of attraction is greater than the grid size used to discretize the function.
4. Finding the global minimum in a fixed time  $T$  is certain if the minimum grid size that can be achieved in that time is less than or equal to the size of the basin of attraction.





## 4 GMG ASSESSMENT ON BENCHMARK FUNCTIONS

The performance of global optimization algorithms is traditionally benchmarked by comparing the number of function evaluations required to find the global minimum, for a standard set of objective functions.

We have assessed GMG’s performance versus that of the TRUST algorithm [12], for seven standard benchmark functions, which are defined in Table 1. The TRUST algorithm uses non-Lipschitzian terminal repellers and subenergy tunneling to provide a deterministic global optimization algorithm that, on the standard benchmark functions, proved significantly faster and more accurate than other previously reported global optimization techniques.

The number of function evaluations reported for TRUST is an average over the complete set of initial starting points, which consists of all the corners of the domain, for locating one global minimum to approximately 4-digit accuracy. The total number of function evaluations for GMG to find one global minimum is the sum of the minimum number of function evaluations that consistently identifies the basin of attraction of the global minimum and the number of iterations required to descend to this global minimum to approximately 4-digit accuracy, using a steepest descent algorithm.

For the Rastrigin function [13], we note that the minimum number of function evaluations depends critically on the symmetry of the problem. This global minimum is co-located with a grid point of a relatively sparse covering grid, resulting in a very small number of required function evaluations, and is a consequence of the highly symmetric domain of this benchmark function. If the global minimum were not located at the origin, or if the domain boundaries were not symmetric with respect to the origin, then the grid density for the GMG search would be greatly increased (e.g., to at least  $12 \times 12$ ), and the GMG algorithm would perform significantly worse than TRUST.

For the Hartman function, the steepest descent algorithm converges very slowly for one of the dimensions. Finding the global minimum by starting the TRUST algorithm from the point identified by GMG requires far fewer function evaluations, so this method was substituted for steepest descent when calculating the total number of function evaluations.

The results in Table 2 show that in four out of seven cases, GMG finds the global minimum using fewer total function evaluations than TRUST. On the other had in the remaining three cases GMG requires a significantly larger number of function evaluations than TRUST. The results of this comparison, which was performed mostly for completeness’s sake, cannot be interpreted in a conclusive manner. Indeed, GMG is fundamentally different from many widely used approaches, including TRUST, in that it uses available additional information in the intrinsic design of the algorithm, provides an *a priori* estimate of the effort needed to obtain the solution, and *guarantees* finding the global minimum. Knowledge of the value of the global minimum may be used by TRUST and other global optimization algorithms to provide a stopping criterium for benchmarks and other informed cases, but is otherwise irresponsible to predict whether these other algorithms will find the global minimum in a finite time.

We note other differences as well. First, some of the benchmark objective functions are degenerate, i.e., they have multiple (equal) global minima. Our approach has been developed to solve parameter identification problems with unique solutions, but it can be easily adapted to find several global minima. This is typically achieved with only a small increase in the required number of function evaluations.

As the dimensionality of the objective function grows, the number of function evaluations

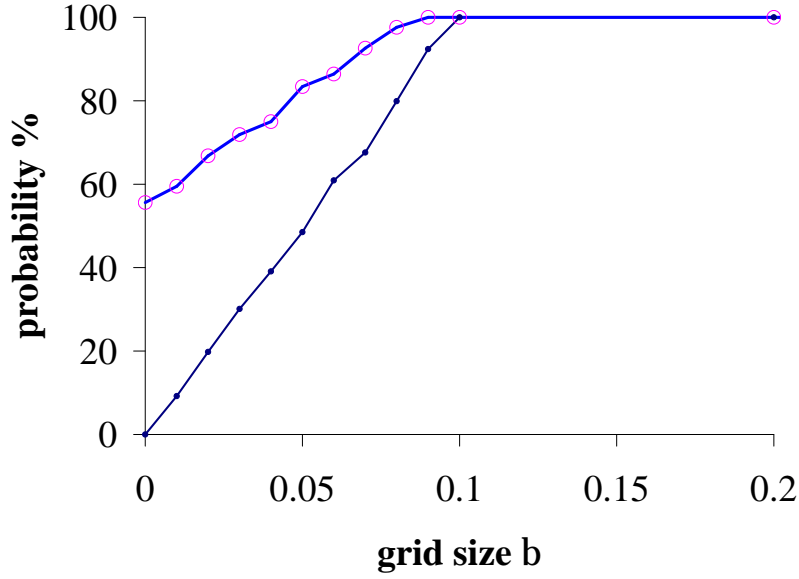


Figure 4: Probability to find the global minimum vs. the grid size for exactly-known ( $\beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.2$ ; hollow dots) one-dimensional golf-course functions.

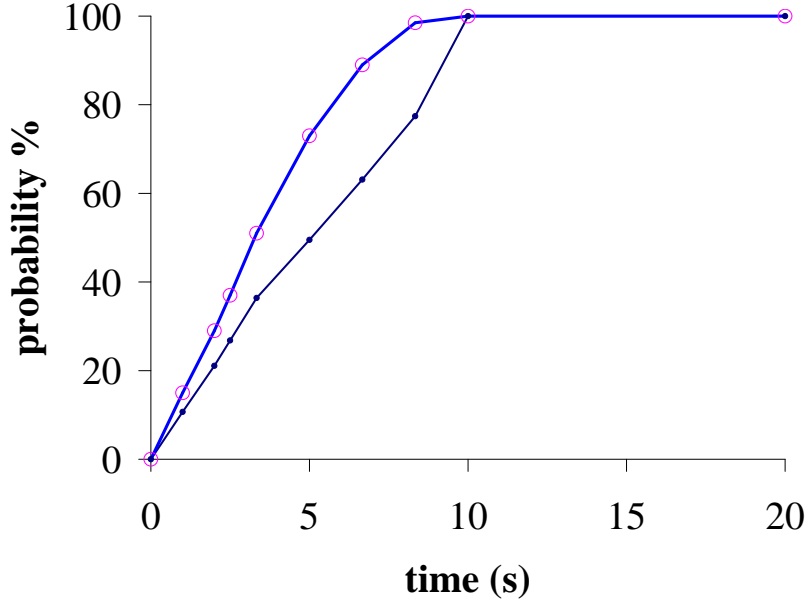


Figure 5: Probability to find the global minimum vs. a fixed running time, for exactly-known ( $\beta_0 = 0.1$ ; solid dots) and “uncertain” ( $\beta_1 = 0.1$ ,  $\beta_2 = 0.2$ ; hollow dots) one-dimensional golf-course functions.

Table 1: Benchmark objective functions.

Function	Definition of objective function
Branin	$f(x_1, x_2) = (x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6)^2 + (10 - \frac{5}{4\pi})\cos[x_1] + 10$
Camelback	$f(x_1, x_2) = x_1^2(4 - 2.1x_1^2 + \frac{x_1^4}{3}) + x_1x_2 + 4x_2^2(x_2^2 - 1)$
Goldstein-Price	$f(x_1, x_2) = \{1 + (1 + x_1 + x_2)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)\} \times \{30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)\}$
Rastrigin	$f(x_1, x_2) = x_1^2 + x_2^2 - \cos[18x_1] - \cos[18x_2]$
Shubert	$f(x_1, x_2) = \left\{ \sum_{i=1}^5 i \cos[(i+1)x_1 + i] \right\} \times \left\{ \sum_{i=1}^5 i \cos[(i+1)x_2 + i] \right\}$
Hartman	$f(x_1, x_2, x_3) = - \sum_{i=1}^4 C_i \exp \left\{ - \sum_{j=1}^3 A_{ij}(x_j - B_{ij}) \right\},$ $A = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix} \quad B = \begin{bmatrix} 0.36890 & 0.1170 & 0.2673 \\ 0.46990 & 0.4387 & 0.7470 \\ 0.10910 & 0.8732 & 0.5547 \\ 0.03815 & 0.5743 & 0.8828 \end{bmatrix} \quad C = \begin{bmatrix} 1.0 \\ 1.2 \\ 3.0 \\ 3.2 \end{bmatrix}$
Styblinski	$f(x_1, \dots, x_5) = \frac{x_1^4 - 16x_1^2 + 5x_1}{2} + \frac{x_2^4 - 16x_2^2 + 5x_2}{2} + \sum_{i=3}^5 (x_i - 1)^2$

Table 2: Benchmark performance comparison.

GOP objective function (# dimensions)	Number of global minima	TRUST average # of fn evaluations	GMG total # of fn evaluations
Branin (2)	3	55	46
Camelback (2)	2	31	30
Goldstein- Price (2)	1	103	69
Rastrigin (2)	1	59	51
Shubert (2)	18	72	2307
Hartman (3)	1	58	104
Styblinski (5)	1	89	1068

for GMG increases exponentially. This is a manifestation of the dimensionality curse, since the size of the grid grows exponentially with the number of dimensions. While this feature of GMG is expected to be general, the number of evaluations required by TRUST with respect to the number of dimensions is difficult to quantify precisely. When sufficient additional information is available, finding the global minimum using GMG is guaranteed. This is something that most global optimization algorithms used in practise cannot achieve in principle.

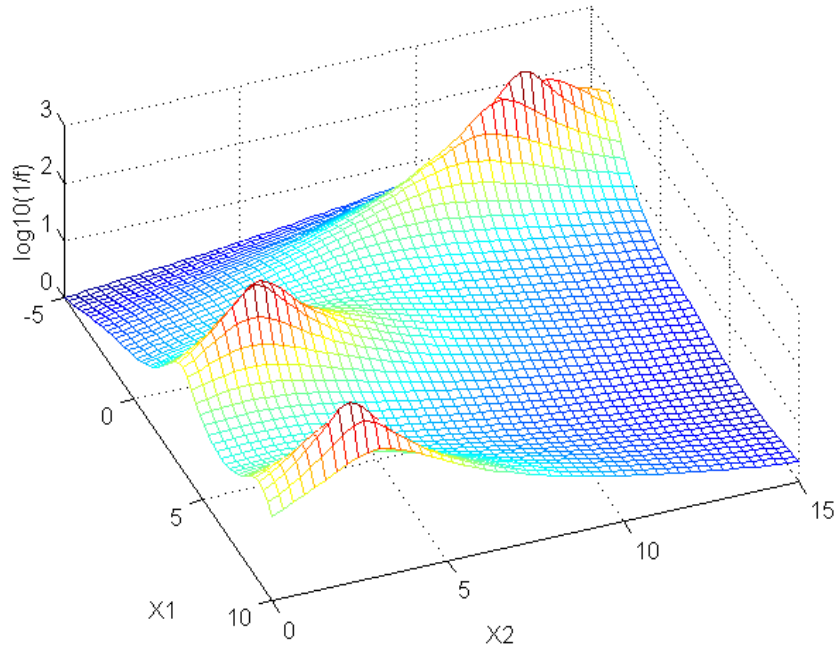


Figure 6: The Branin objective function.

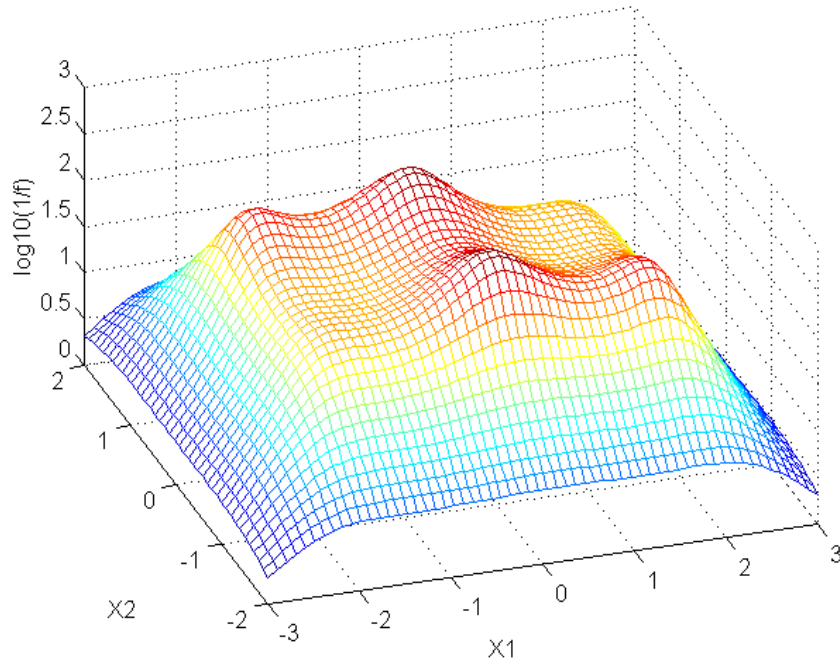


Figure 7: The Camelback objective function.

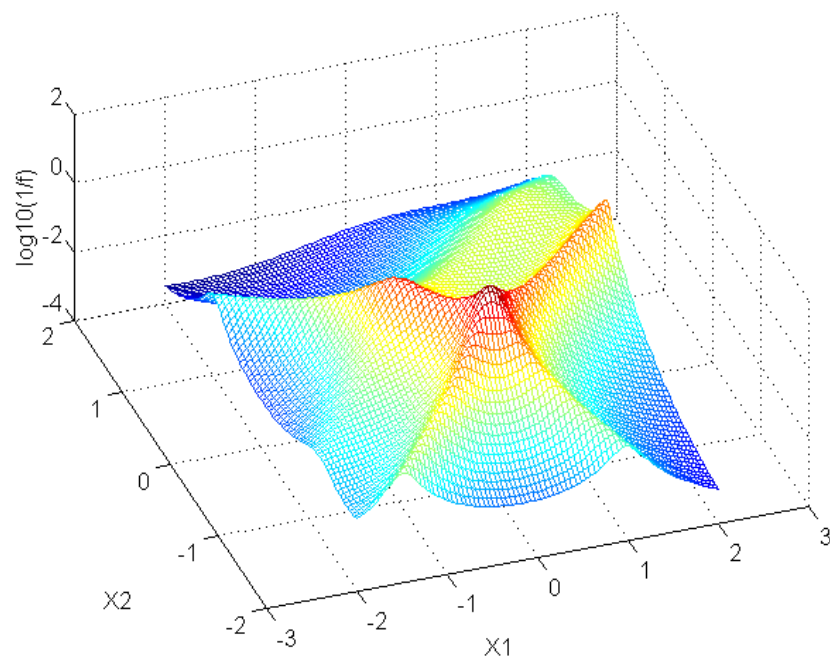


Figure 8: The Goldstein-Price objective function.

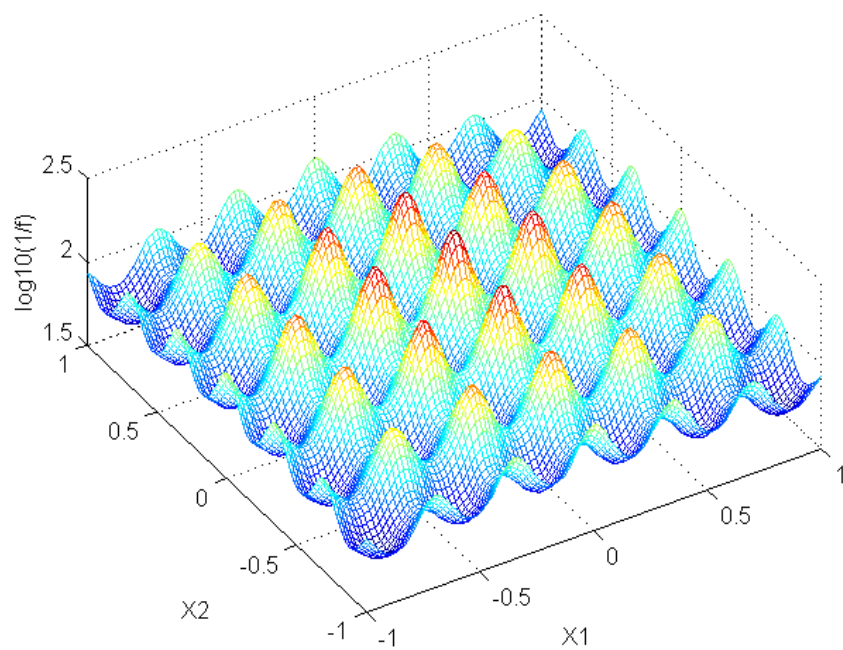


Figure 9: The Rastrigin objective function.

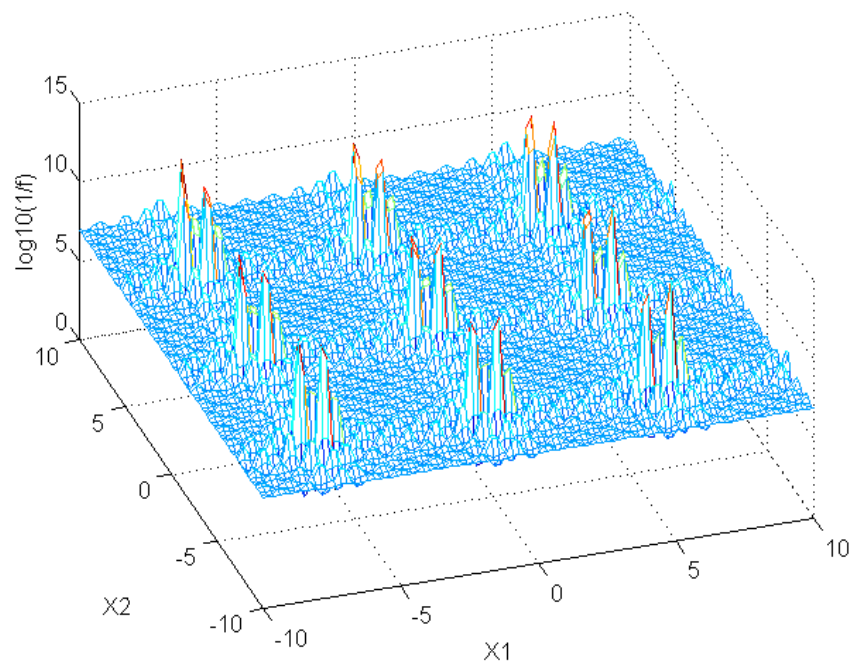


Figure 10: The Shubert objective function.

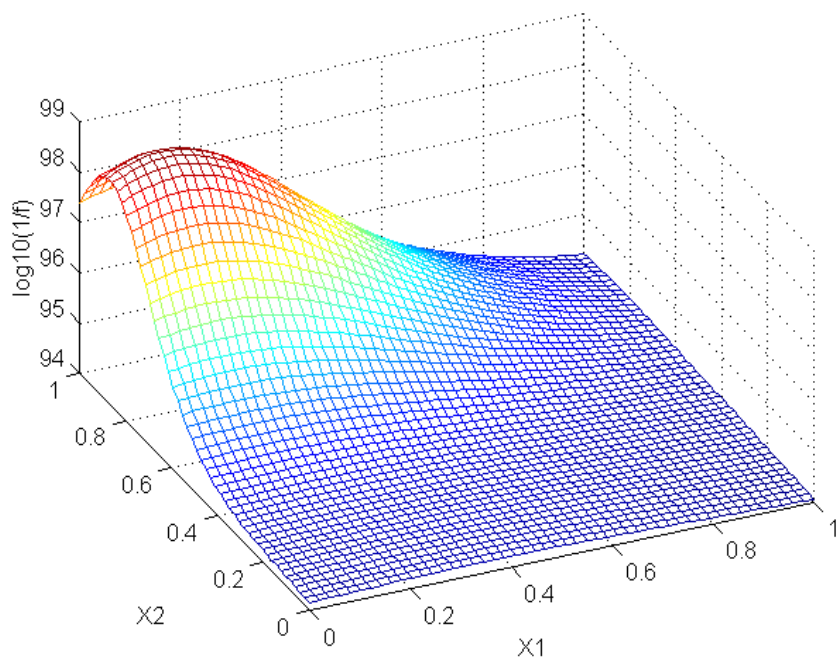


Figure 11: The Hartman objective function.





## 5 SOLVING THE MPR PROBLEM

The first challenge in implementing GMG for the MPR problem, was to generate the error function,  $E(\mathbf{y})$ . The standard mean profile  $\rho_0(r)$  used throughout this paper was taken from high-quality measurements by NASA [16] for mid-latitudes in summer, and is defined from the ground to 50 km, as shown in Fig. 12.

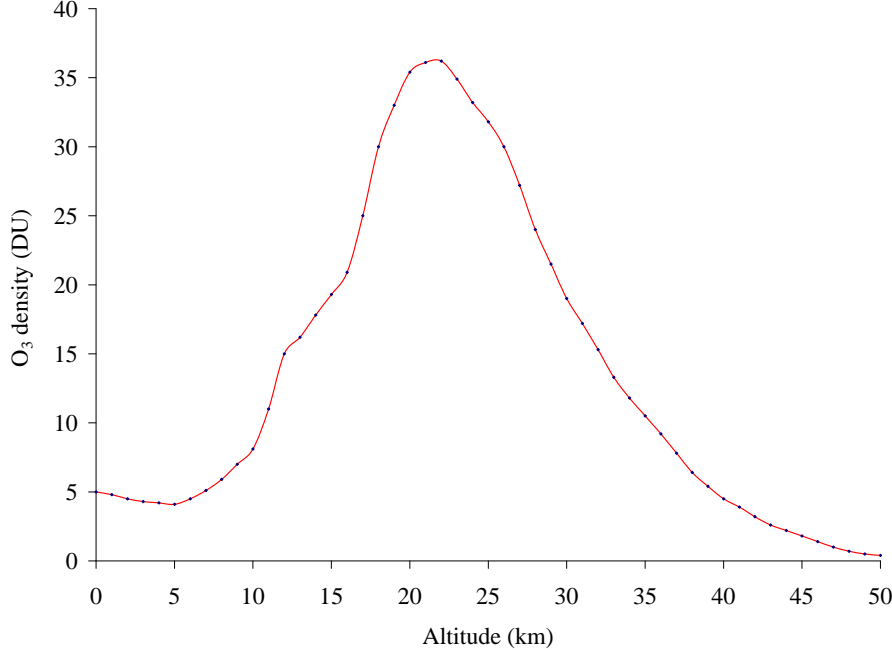
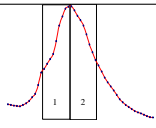
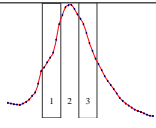
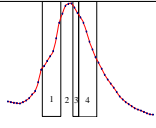
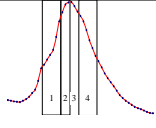
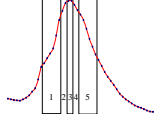


Figure 12: Standard mean ozone profile for mid-latitudes in summer. The ozone density unit used in the graph is the Dobson Unit ( $1 \text{ DU} = 2.14 \times 10^{-8} \text{ kg/m}^3$ ).

In order to resolve salient features of the ozone profile, MODTRAN needs a resolution of 1 km for the standard mean ozone profile. However, we do not have the computational capability to search for variations of the ozone profile among 50 dimensions (1-50 km), since the cost of evaluating even one value of the error function using MODTRAN is relatively high. Hence, only a small number ( $k = 3, 4, 5$ ) of variations in the ozone profile,  $\delta\rho_1, \delta\rho_2, \dots, \delta\rho_k$ , were used in the evaluation of the error function. Each variation acts independently on the ozone profile, across an atmospheric layer covering specific altitudes. The magnitude of the variations is restricted to within  $\pm 1\%$  of the corresponding standard mean ozone profile. The partitioning of the atmospheric layers has a significant impact on the target radiance, as discussed below (see Table 3).

As mentioned in Section 2, MODTRAN was used to calculate a proxy for the single sensor measurement,  $S$ , of the radiance from a target at a known range, for a specific ozone profile. The target range,  $x_0$ , was chosen at random from values between 120 and 130 km, corresponding to typical baseline MPR scenarios for an airborne sensor. The variations in the ozone profile,  $\delta\rho_{iS}(r)$ , were also chosen randomly.

Table 3: Percentage error in the target radiance vs. the number of ozone layers.

# O3 layers	Layer 1 (km)	Layer 2 (km)	Layer 3 (km)	Layer 4 (km)	Layer 5 (km)	Layer partition	Deviation (%)
2	13-22	22-30	-	-	-		$41.6 \pm 0.1$
3	13-19	19-25	25-30	-	-		$(1.7 \pm 1.1) \times 10^{-3}$
4	13-19	19-23	23-25	25-30	-		$(1.4 \pm 1.1) \times 10^{-3}$
4	13-19	19-22	22-25	25-30	-		$6.79 \pm 0.03$
5	13-19	19-21	21-23	23-25	25-30		0.0

Next, the radiance  $R(x, \delta\rho_1, \delta\rho_2, \dots, \delta\rho_k)$  at the sensor was calculated using MODTRAN, for ranges in the interval 120-130 km, and for different variations in the ozone profile, within  $\pm 1\%$  of  $\rho_0(r)$ . Both the predicted radiance and the measurement  $S$  were calculated by summing up the spectral radiances across the same ozone-absorption window (9-11  $\mu\text{m}$ ), which does not suffer much from water absorption.

When implementing the algorithm, we gradually increased the number of ozone layers from two to five. The thickness of the layers varied between 2 and 8 km, as shown in Table 3. This provided a general outlook for the dependence of the target radiance on the partitioning of the ozone profile. In particular, it showed that the radiance is most sensitive to the resolution of the layer with the highest density of ozone. Since detection of airborne missiles with remote sensing aircraft [4] uses an uplooking baseline, MPR is particularly sensitive to the peak ozone density at  $\approx 22$  km. Therefore, to obtain the best sensitivity available with a small number of ozone layers in our algorithm, the ozone profile should be partitioned nonuniformly, by using more layers at high-density altitudes (19 – 24 km).

Table 3 shows the percentage deviation in the target radiance as compared to the target radiance computed using the highest resolution considered for the ozone concentration (i.e., five layers). This provides a measure of the sensitivity in the construction of the ozone layers as their number is increased from two layers to five layers. In addition, the partitioning or structure of the layers plays an important role in determining the sensitivity. For example,

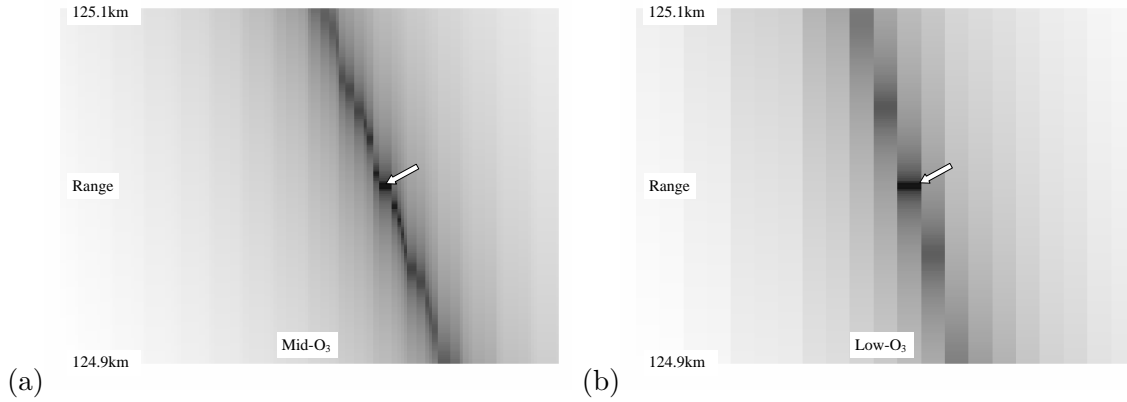


Figure 13: Error function using three ozone layers, for (a) mid-ozone concentration change ( $\pm 1\%$ ) vs. range (km), and b) low-ozone concentration change ( $\pm 1\%$ ) vs. range (km). The error function is plotted on a logarithmic scale, with dark intensities indicating smaller values. The position of the global minimum is indicated by the superimposed arrows.

if the peak ozone density (21 – 22 km) is designed to reside within the interior of one layer, the deviation between the radiance of the 4-layer and the 5-layer models is negligible:  $(1.4 \pm 1.1) \times 10^{-3}\%$ . However, the deviation increases greatly if the 4-layer model places the peak ozone densities on the boundary between two adjacent layers.

The resulting error function contained tens ( $k = 3$ ) to thousands ( $k = 5$ ) of values close to zero, thus providing a good testing environment for our GOP algorithm. However, these values are not separate minima, but rather belong to the same basin of attraction. Fig. 13 shows the intensity plots of slices of an error function obtained using three ozone layers. The ozone profile has been partitioned uniformly into a low layer (13 – 18 km), a mid layer (19 – 24 km) and a high layer (25 – 30 km). The sensor is assumed to be at a typical altitude of 11 km, and the target is at an altitude of 33 km, as illustrated for the upward-looking scenario in Fig. 1.

The range interval shown in Fig. 13 focuses on the region that contains the target range, which is set to 125 km. The values of the error function for ranges within  $\pm 0.1$  km are very close to the global minimum. We note that a small uncertainty in the variation of the mid-ozone layer can result in a large uncertainty ( $\approx \pm 0.1$  km) for the target range. Thus solving the GOP for the MPR problem yields the target range with an accuracy that is better by at least an order of magnitude, compared to algorithms that ignore the variation in the mid-ozone layer.

GMG was implemented in Fortran 90 to simplify calling MODTRAN (Fortran 77 code) when evaluating the error function. The algorithm takes full advantage of the IBM Power4 parallel supercomputer within ORNL's Center for Computational Sciences, which hosts a total of  $27 \times 32$  1.3GHz processors. The most demanding computation required 441 processors,  $\approx 55\%$  of the total available, which enabled us to use six dimensions for the radiance, i.e., five ozone layers and the range. In this case, the typical run time was of the order of several hours. The rather voluminous output that is normally produced by MODTRAN was trimmed to a few hundred megabytes - however the output of our algorithm, describing the error function and the position of the global minimum was much smaller.

## 5.1 GMG ASSESSMENT ON THE MPR PROBLEM

In order to assess the performance of GMG for the MPR problem, we first verified it in the ideal case, where complete and unadulterated additional information was known about the error function, as described in Section 3.1. The basin of attraction of the global minimum was found in 100% of trials. After descending to the global minimum (if required), the range  $x_0$  and the ozone densities  $\rho_1, \rho_2, \dots, \rho_k$  found by the algorithm were *identical* to the values used originally to calculate the sensor measurement  $S$ .

The performance and time scaling of the algorithm are consistent with the previous results shown for the golf-course problem, namely:

1. The total processor time to find a solution grows exponentially with the number of dimensions. This is compounded by computationally intensive calls to MODTRAN, and limits the number of dimensions that can be modelled.
2. For a fixed number of dimensions, the total processor time grows as an inverse power of the size of the basin of attraction. As expected, the algorithm is most computationally demanding when the size of the basin of attraction of the global minimum is relatively small.
3. The probability of finding the global minimum is equal to one if the minimum size of the basin of attraction is greater than the grid size used to discretize the function.
4. The probability to find the global minimum in a fixed time  $T$  is equal to one if the minimum grid size that can be achieved in that time is less than or equal to the size of the basin of attraction.

The amount of information known about the error function was then relaxed to simulate uncertainties typically present in MPR. In particular, the information related to the size of the basin of attraction for the global minimum is of prime importance for an efficient guarantee.

Taking the size of the basin of attraction as unknown, we progressively increased the grid density by a factor of 2, to search for a unique solution that satisfied the threshold in the discrete transformation in Eq. 4. This approach is guaranteed to find the basin of attraction of the global minimum, if the value of the global minimum is zero (or some other known value), or if the threshold  $\delta$  is known. However, we lose the guarantee if no information of types (i) or (ii) is available, due to noise, etc.

By implementing GMG to various realizations of the MPR problem, we also verified that partitioning the ozone profile has a significant effect on the target radiance. In particular, we found that the error function depends very sensitively on the mid-ozone layer, containing the peak ozone density. As a result, the basin of attraction of the global minimum is typically localized in the direction of this variable (see Fig. 13 for two-dimensional cuts of the four-dimensional error function). If this additional information about the error function is used, e.g., the peak ozone density is (approximately) identified beforehand, then we can optimize the use of computational resources by focusing from the beginning on those layers that cause the greatest uncertainty in the error function. Incidentally, this would provide another instance in which additional information leads to a considerable reduction of the computational complexity.

## 6 CONCLUSION

We developed and validated GMG, a new deterministic algorithm for GOP's, which guarantees to find the global minimum. The novel idea behind GMG is to identify and systematically use additional information about the objective function. Unlike previous similar approaches, the additional information we use is not directly related to smoothness properties. Our algorithm reduces the computational complexity of the problem, and thereby guarantees finding the location of the global minimum in a reasonably short time. Just as importantly, GMG allows us to predict an upper bound for the computational effort required to find the global minimum.

After validating GMG on various benchmark problems, we applied it to the MPR problem and assessed its performance, which turned out to be excellent. This demonstrates that GMG can be applied to GOP's related to national security, computational biology, materials science etc., which often require precise, guaranteed and timely identification of the global minimum.



## References

- [1] C. A. Floudas and P. M. Pardalos eds., *State of the Art in Global Optimization: Computational Methods and Applications*, (Dordrecht: Kluwer, 1996).
- [2] A. Törn and A. Žilinkas, *Global Optimization*, (Berlin: Springer-Verlag, 1989).
- [3] R. Horst and H. Tuy, *Global Optimization*, 2nd ed., (Berlin: Springer-Verlag, 1993).
- [4] G. Scriven, N. Gat, and R. Lyons, *Monocular Passive Ranging Sensitivity Analysis and Error Minimization*, AFRL Technical Report, 2001.
- [5] W. Jeffrey, J. S. Draper, and R. Gobel, *Monocular Passive Ranging*, IRIS Targets, Backgrounds and Discrimination Conference (1994).
- [6] D. Kleppner, F. K. Lamb, and D. E. Mosher, *Physics Today* **57**, 30-35 (January 2004).
- [7] A. Berk, G. P. Anderson, L. S. Bernstein, P. K. Acharya, H. Dothe, M. W. Matthew, S. M. Adler-Golden, J. H. Chetwynd, Jr., S. C. Richtsmeier, B. Pukall, C. L. Allred, L. S. Jeong, and M. L. Hoke *MODTRAN4 Radiative Transfer Modeling for Atmospheric Correction*, SPIE Proceeding, Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, Volume 3756, July 1999.
- [8] P. K. Acharya, A. Berk, G. P. Anderson, N. F. Larsen, S.-Chee Tsay, and K. H. Stamnes, *MODTRAN4: Multiple Scattering and Bi-Directional Reflectance Distribution Function (BRDF) Upgrades to MODTRAN*, SPIE Proceeding, Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, Volume 3756, July 1999.
- [9] ADIFOR 2.0: Automatic Differentiation of Fortran; ([www-fp.mcs.anl.gov/autodiff/](http://www-fp.mcs.anl.gov/autodiff/)).
- [10] V. Protopopescu and J. Barhen, *Phys. Lett. A* **296**, 9 (2002).
- [11] C. D’Helon and V. Protopopescu, *J. Phys. A* **35**, L597 (2002).
- [12] J. Barhen, V. Protopopescu, and D. Reister, *Science* **276**, 1094 (1997).
- [13] R. Horst and P. M. Pardalos, eds., *Handbook of Global Optimization*, (Dordrecht: Kluwer, 1995).
- [14] A. Berk, L. S. Bernstein, G. P. Anderson, P. K. Acharya, D. C. Robertson, J. H. Chetwynd, and S. M. Adler-Golden, *MODTRAN Cloud and Multiple Scattering Upgrades with application to AVIRIS*, Remote Sens. Environ., **65**, 367 (1998).
- [15] In general, the basin of attraction of the global minimum is multidimensional. If the extent of the basin of attraction is known for each dimension, then the resolution of the grid can be set accordingly for each dimension.
- [16] Studying Earth’s Environment from Space (SEES): ([www.ccpo.odu.edu/SEES/index.html](http://www.ccpo.odu.edu/SEES/index.html)).





### **Internal Distribution**

- |                                  |                                    |
|----------------------------------|------------------------------------|
| 1-5. Jacob Barhen                | 18. ORNL Laboratory Records - OSTI |
| 6-10. Cassius D'Helon            | 19-23. Vladimir Protopopescu       |
| 11. Central Research Library     | 24. David Reister                  |
| 12. Markus Eisenbach             | 25. Thomas Schulthess              |
| 13. Linda Horton                 | 26. Michael Leon Simpson           |
| 14. Neena Imam                   | 27. Terrence P. Sjoreen            |
| 15. Suzanne Lenhart              | 28. Malcolm Stocks                 |
| 16. Jeff Nichols                 | 29-33. Jack C. Wells               |
| 17. ORNL Laboratory Records - RC |                                    |

### **External Distribution**

1. Alberto Elfes, Strategic Systems Technology Program Office, Jet Propulsion Laboratory, 4800 Oak Grove Drive, MS-179-224, Pasadena, CA 91109-8099 [elfes@jpl.nasa.gov](mailto:elfes@jpl.nasa.gov)
2. Charles Weisbin, NASA Jet Propulsion Laboratory, Strategic Systems Technology Program Office, Mail Code 179-224, Pasadena, CA 91129 [Charles.R.Weisbin@jpl.nasa.gov](mailto:Charles.R.Weisbin@jpl.nasa.gov)
3. Timothy J. Fitzsimmons, SC-13 / Germantown Building, U. S. Department of Energy, 1000 Independence Ave., S. W., Washington, DC 20585-1290  
[Tim.Fitzsimmons@science.doe.gov](mailto:Tim.Fitzsimmons@science.doe.gov)